

RECEIVED

JUN 25 2001

Page 1 of 7

0590  
1641

#6



TECH CENTER 1600/2900

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/733,757

DATE: 06/01/2001

TIME: 11:16:56

Input Set : A:\es.txt

Output Set: C:\CRF3\06012001\I733757.raw

P.S.

3 <110> APPLICANT: Mack, David  
4 Gish, Kurt  
5 Wilson, Keith  
7 <120> TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS  
AND METHODS

8 OF SCREENING FOR COLORECTAL CANCER MODULATORS

10 &lt;130&gt; FILE REFERENCE: A-69796/DJB/JJD

12 &lt;140&gt; CURRENT APPLICATION NUMBER: US 09/733,757

13 &lt;141&gt; CURRENT FILING DATE: 2000-12-08

15 &lt;160&gt; NUMBER OF SEQ ID NOS: 3

17 &lt;170&gt; SOFTWARE: PatentIn version 3.0

19 &lt;210&gt; SEQ ID NO: 1

20 &lt;211&gt; LENGTH: 2103

21 &lt;212&gt; TYPE: DNA

22 &lt;213&gt; ORGANISM: Homo sapiens

24 &lt;220&gt; FEATURE:

25 &lt;221&gt; NAME/KEY: CDS

26 &lt;222&gt; LOCATION: (70)..(2103)

28 &lt;400&gt; SEQUENCE: 1

ENTERED

29 cggcagcagg cgcgactgcg aggcctggacg ytacgggctc ctggaaagga gacaccagca 60  
31 ttgtccaca atg ctg tca tcc act gac ttt aca ttt gct tcc tgg gag ctt 111  
32 Met Leu Ser Ser Thr Asp Phe Thr Phe Ala Ser Trp Glu Leu  
33 1 5 10  
35 gtg gtc cgc gtt gac cat ccc aat gaa gag cag cag aaa gac gtc aca 159  
36 Val Val Arg Val Asp His Pro Asn Glu Glu Gln Gln Lys Asp Val Thr  
37 15 20 25 30  
39 ctg aga gta tct gga gac ctt cac gtt gga gga gtg atg ctc aag tta 207  
40 Leu Arg Val Ser Gly Asp Leu His Val Gly Gly Val Met Leu Lys Leu  
41 35 40 45  
43 gta gaa cag atc aat ata tcc caa gac tgg tca gac ttt gct ctt tgg 255  
44 Val Glu Gln Ile Asn Ile Ser Gln Asp Trp Ser Asp Phe Ala Leu Trp  
45 50 55 60  
47 tgg gaa cag aag cat tgc tgg ctt ctg aaa acc cac tgg acc ctg gac 303  
48 Trp Glu Gln Lys His Cys Trp Leu Leu Lys Thr His Trp Thr Leu Asp  
49 65 70 75  
51 aaa tat ggg gtc cag gca gat gca aag ctt ctc ttc acc cct cag cat 351  
52 Lys Tyr Gly Val Gln Ala Asp Ala Lys Leu Leu Phe Thr Pro Gln His  
53 80 85 90  
55 aaa atg ctg cgc ctt cgt ctg ccg aat ttg aag atg gtg agg ttg cga 399  
56 Lys Met Leu Arg Leu Arg Leu Pro Asn Leu Lys Met Val Arg Leu Arg  
57 95 100 105 110  
59 gtc agc ttc tca gct gtg gtt ttt aaa gct gtc agt gat atc tgc aaa 447  
60 Val Ser Phe Ser Ala Val Val Phe Lys Ala Val Ser Asp Ile Cys Lys  
61 115 120 125  
63 atc ctg aat att aga aga tca gaa gag ctt tcc ttg tta aag ccg tct 495  
64 Ile Leu Asn Ile Arg Arg Ser Glu Glu Leu Ser Leu Leu Lys Pro Ser  
65 130 135 140  
67 ggt gac tat ttt aag aag aag aag aaa aaa gac aaa aat aat aag gaa 543

## RAW SEQUENCE LISTING

DATE: 06/01/2001

PATENT APPLICATION: US/09/733,757

TIME: 11:16:56

Input Set : A:\es.txt

Output Set: C:\CRF3\06012001\I733757.raw

```

68 Gly Asp Tyr Phe Lys Lys Lys Lys Lys Lys Asp Lys Asn Asn Lys Glu
69      145      150      155
71 ccc ata att gaa gat att cta aac ctg gag agt tct cca aca gct tca      591
72 Pro Ile Ile Glu Asp Ile Leu Asn Leu Glu Ser Ser Pro Thr Ala Ser
73      160      165      170
75 ggt tca tca gta agt cct ggt tta tac agt aaa acc atg acc cct ata      639
76 Gly Ser Ser Val Ser Pro Gly Leu Tyr Ser Lys Thr Met Thr Pro Ile
77 175      180      185      190
79 tat gac ccc atc aat gga aca cca gca tca tcc acc atg act tgg ttc      687
80 Tyr Asp Pro Ile Asn Gly Thr Pro Ala Ser Ser Thr Met Thr Trp Phe
81      195      200      205
83 agt gac agc cct ttg acg gaa caa aac tgc agc atc ctc gca ttc agc      735
84 Ser Asp Ser Pro Leu Thr Glu Gln Asn Cys Ser Ile Leu Ala Phe Ser
85      210      215      220
87 caa ccc ccc cag tcc cca gaa gca ctt gcg gat atg tac cag cct cgg      783
88 Gln Pro Pro Gln Ser Pro Glu Ala Leu Ala Asp Met Tyr Gln Pro Arg
89      225      230      235
91 tct ctg gtt gat aaa gcc aag ctc aat gca ggt tgg cta gac tcc tca      831
92 Ser Leu Val Asp Lys Ala Lys Leu Asn Ala Gly Trp Leu Asp Ser Ser
93      240      245      250
95 cgc tcc ctt atg gaa caa ggc atc caa gag gat gag cag ctg ctc tta      879
96 Arg Ser Leu Met Glu Gln Gly Ile Gln Glu Asp Glu Gln Leu Leu
97 255      260      265      270
99 cga ttt aaa tat tat tct ttc ttc gac ttg aat cct aaa tat gat gct      927
100 Arg Phe Lys Tyr Tyr Ser Phe Phe Asp Leu Asn Pro Lys Tyr Asp Ala
101      275      280      285
103 gtc cga ata aac caa ctc tat gag caa gcc agg tgg gcc att ctc tta      975
104 Val Arg Ile Asn Gln Leu Tyr Glu Gln Ala Arg Trp Ala Ile Leu Leu
105      290      295      300
107 gaa gaa att gat tgc aca gag gaa gaa atg ttg atc ttt gca gct cta      1023
108 Glu Glu Ile Asp Cys Thr Glu Glu Glu Met Leu Ile Phe Ala Ala Leu
109      305      310      315
111 cag tac cac att agc aaa ctg tcg ttg tct gct gaa aca cag gat ttt      1071
112 Gln Tyr His Ile Ser Lys Leu Ser Leu Ser Ala Glu Thr Gln Asp Phe
113      320      325      330
115 gca ggc gag tcc gag gtt gat gaa ata gaa gcg gcg ctt tct aat ttg      1119
116 Ala Gly Glu Ser Glu Val Asp Glu Ile Glu Ala Ala Leu Ser Asn Leu
117 335      340      345      350
119 gaa gta acc cta gaa ggt gga aaa gcg gac agc ctt ttg gag gac att      1167
120 Glu Val Thr Leu Glu Gly Gly Lys Ala Asp Ser Leu Leu Glu Asp Ile
121      355      360      365
123 act gat atc cct aaa ctt gca gat aat ctc aaa tta ttt agg ccc aag      1215
124 Thr Asp Ile Pro Lys Leu Ala Asp Asn Leu Lys Leu Phe Arg Pro Lys
125      370      375      380
127 aag tta cta cca aaa gct ttc aaa caa tat tgg ttt atc ttt aaa gac      1263
128 Lys Leu Leu Pro Lys Ala Phe Lys Gln Tyr Trp Phe Ile Phe Lys Asp
129      385      390      395
131 aca tcc ata gca tac ttt aaa aat aag gaa ctt gaa caa gga gaa cca      1311
132 Thr Ser Ile Ala Tyr Phe Lys Asn Lys Glu Leu Glu Gln Gly Glu Pro

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/733,757

DATE: 06/01/2001

TIME: 11:16:57

Input Set : A:\es.txt

Output Set: C:\CRF3\06012001\I733757.raw

133	400	405	410	
135	cta gaa aaa cta aat ctt	aga ggc tgc gaa gtt	gtg ccc gat gta aat	1359
136	Leu Glu Lys Leu Asn Leu	Arg Gly Cys Glu Val	Val Pro Asp Val Asn	
137	415	420	425	430
139	gta gca gga aga aaa ttt	gga atc aag tta cta	atc cct gtt gcc gat	1407
140	Val Ala Gly Arg Lys Phe	Gly Ile Lys Leu Leu	Ile Pro Val Ala Asp	
141	435	440	445	
143	ggt atg aat gaa atg tat	ttg aga tgt gac cat	gag aat caa tac gcc	1455
144	Gly Met Asn Glu Met Tyr	Leu Arg Cys Asp His	Glu Asn Gln Tyr Ala	
145	450	455	460	
147	caa tgg atg gct gcc tgc	atg ttg gca tcg aag	ggc aaa acc atg gca	1503
148	Gln Trp Met Ala Ala Cys	Met Leu Ala Ser Lys	Gly Lys Thr Met Ala	
149	465	470	475	
151	gac agc tcc tac cag cca	gag gtc ctc aac atc	ctt tca ttt ctg agg	1551
152	Asp Ser Ser Tyr Gln Pro	Glu Val Leu Asn Ile	Leu Ser Phe Leu Arg	
153	480	485	490	
155	atg aaa aac agg aac tct	gca tct cag gtg gct	tcc agt ctc gaa aac	1599
156	Met Lys Asn Arg Asn Ser	Ala Ser Gln Val Ala	Ser Ser Leu Glu Asn	
157	495	500	505	510
159	atg gat atg aac cca gaa	tgt ttt gtg tca cca	cgg tgt gca aag aaa	1647
160	Met Asp Met Asn Pro Glu	Cys Phe Val Ser Pro	Arg Cys Ala Lys Lys	
161	515	520	525	
163	cac aaa tcc aaa cag ctg	gcc gcc cgg atc ctg	gag gcg cac cag aac	1695
164	His Lys Ser Lys Gln Leu	Ala Ala Arg Ile Leu	Glu Ala His Gln Asn	
165	530	535	540	
167	gtg gcc cag atg ccc ctg	gtc gaa gcc aag ctg	cgg ttc atc cag gcg	1743
168	Val Ala Gln Met Pro Leu	Val Glu Ala Lys Leu	Arg Phe Ile Gln Ala	
169	545	550	555	
171	tgg cag tca ctg cct gag	ttt ggc ctc acc tac	tac ctt gtc aga ttt	1791
172	Trp Gln Ser Leu Pro Glu	Phe Gly Leu Thr Tyr	Tyr Leu Val Arg Phe	
173	560	565	570	
175	aaa gga agc aaa aaa gat	gac att ctg gga gtt	tca tat aac agg ttg	1839
176	Lys Gly Ser Lys Lys Asp	Asp Ile Leu Gly Val	Ser Tyr Asn Arg Leu	
177	575	580	585	590
179	att aaa att gat gca gcc	acc ggg att cca gtg	aca aca tgg aga ttc	1887
180	Ile Lys Ile Asp Ala Ala	Thr Gly Ile Pro Val	Thr Thr Trp Arg Phe	
181	595	600	605	
183	aca aat atc aaa cag tgg	aat gta aac tgg gaa	acc cgg cag gtg gtc	1935
184	Thr Asn Ile Lys Gln Trp	Asn Val Asn Trp Glu	Thr Arg Gln Val Val	
185	610	615	620	
187	atc gag ttt gag caa aac	gtc ttt act gct ttc	acc tgc ctg agt gca	1983
188	Ile Glu Phe Asp Gln Asn	Val Phe Thr Ala Phe	Thr Cys Leu Ser Ala	
189	625	630	635	
191	gat tgc aag att gtg cac	gag tac att ggc ggc	tac att ttc ttg tcc	2031
192	Asp Cys Lys Ile Val His	Glu Tyr Ile Gly Gly	Tyr Ile Phe Leu Ser	
193	640	645	650	
195	acc cgc tcc aag gac cag	aat gaa aca ctc gat	gag gac ttg ttc cac	2079
196	Thr Arg Ser Lys Asp Gln	Asn Glu Thr Leu Asp	Glu Asp Leu Phe His	
197	655	660	665	670

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/733,757

DATE: 06/01/2001

TIME: 11:16:57

Input Set : A:\es.txt

Output Set: C:\CRF3\06012001\I733757.raw

```

199 aaa ttg acc ggc ggt cag gat taa
200 Lys Leu Thr Gly Gly Gln Asp
201                                     675
204 <210> SEQ ID NO: 2
205 <211> LENGTH: 677
206 <212> TYPE: PRT
207 <213> ORGANISM: Homo sapiens
209 <400> SEQUENCE: 2
211 Met Leu Ser Ser Thr Asp Phe Thr Phe Ala Ser Trp Glu Leu Val Val
212 1                                     5 10 15
215 Arg Val Asp His Pro Asn Glu Glu Gln Gln Lys Asp Val Thr Leu Arg
216                                     20 25 30
219 Val Ser Gly Asp Leu His Val Gly Gly Val Met Leu Lys Leu Val Glu
220                                     35 40 45
223 Gln Ile Asn Ile Ser Gln Asp Trp Ser Asp Phe Ala Leu Trp Trp Glu
224 50                                     55 60
227 Gln Lys His Cys Trp Leu Leu Lys Thr His Trp Thr Leu Asp Lys Tyr
228 65                                     70 75 80
231 Gly Val Gln Ala Asp Ala Lys Leu Leu Phe Thr Pro Gln His Lys Met
232                                     85 90 95
235 Leu Arg Leu Arg Leu Pro Asn Leu Lys Met Val Arg Leu Arg Val Ser
236                                     100 105 110
239 Phe Ser Ala Val Val Phe Lys Ala Val Ser Asp Ile Cys Lys Ile Leu
240                                     115 120 125
243 Asn Ile Arg Arg Ser Glu Glu Leu Ser Leu Leu Lys Pro Ser Gly Asp
244 130                                     135 140
247 Tyr Phe Lys Lys Lys Lys Lys Lys Asp Lys Asn Asn Lys Glu Pro Ile
248 145                                     150 155 160
251 Ile Glu Asp Ile Leu Asn Leu Glu Ser Ser Pro Thr Ala Ser Gly Ser
252                                     165 170 175
255 Ser Val Ser Pro Gly Leu Tyr Ser Lys Thr Met Thr Pro Ile Tyr Asp
256                                     180 185 190
259 Pro Ile Asn Gly Thr Pro Ala Ser Ser Thr Met Thr Trp Phe Ser Asp
260                                     195 200 205
263 Ser Pro Leu Thr Glu Gln Asn Cys Ser Ile Leu Ala Phe Ser Gln Pro
264 210                                     215 220
267 Pro Gln Ser Pro Glu Ala Leu Ala Asp Met Tyr Gln Pro Arg Ser Leu
268 225                                     230 235 240
271 Val Asp Lys Ala Lys Leu Asn Ala Gly Trp Leu Asp Ser Ser Arg Ser
272                                     245 250 255
275 Leu Met Glu Gln Gly Ile Gln Glu Asp Glu Gln Leu Leu Leu Arg Phe
276                                     260 265 270
279 Lys Tyr Tyr Ser Phe Phe Asp Leu Asn Pro Lys Tyr Asp Ala Val Arg
280                                     275 280 285
283 Ile Asn Gln Leu Tyr Glu Gln Ala Arg Trp Ala Ile Leu Leu Glu Glu
284 290                                     295 300
287 Ile Asp Cys Thr Glu Glu Glu Met Leu Ile Phe Ala Ala Leu Gln Tyr
288 305                                     310 315 320
291 His Ile Ser Lys Leu Ser Leu Ser Ala Glu Thr Gln Asp Phe Ala Gly

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/733,757

DATE: 06/01/2001

TIME: 11:16:57

Input Set : A:\es.txt

Output Set: C:\CRF3\06012001\I733757.raw

```

292          325          330          335
295 Glu Ser Glu Val Asp Glu Ile Glu Ala Leu Ser Asn Leu Glu Val
296          340          345          350
299 Thr Leu Glu Gly Gly Lys Ala Asp Ser Leu Leu Glu Asp Ile Thr Asp
300          355          360          365
303 Ile Pro Lys Leu Ala Asp Asn Leu Lys Leu Phe Arg Pro Lys Lys Leu
304          370          375          380
307 Leu Pro Lys Ala Phe Lys Gln Tyr Trp Phe Ile Phe Lys Asp Thr Ser
308 385          390          395          400
311 Ile Ala Tyr Phe Lys Asn Lys Glu Leu Glu Gln Gly Glu Pro Leu Glu
312          405          410          415
315 Lys Leu Asn Leu Arg Gly Cys Glu Val Val Pro Asp Val Asn Val Ala
316          420          425          430
319 Gly Arg Lys Phe Gly Ile Lys Leu Leu Ile Pro Val Ala Asp Gly Met
320          435          440          445
323 Asn Glu Met Tyr Leu Arg Cys Asp His Glu Asn Gln Tyr Ala Gln Trp
324          450          455          460
327 Met Ala Ala Cys Met Leu Ala Ser Lys Gly Lys Thr Met Ala Asp Ser
328 465          470          475          480
331 Ser Tyr Gln Pro Glu Val Leu Asn Ile Leu Ser Phe Leu Arg Met Lys
332          485          490          495
335 Asn Arg Asn Ser Ala Ser Gln Val Ala Ser Ser Leu Glu Asn Met Asp
336          500          505          510
339 Met Asn Pro Glu Cys Phe Val Ser Pro Arg Cys Ala Lys Lys His Lys
340          515          520          525
343 Ser Lys Gln Leu Ala Ala Arg Ile Leu Glu Ala His Gln Asn Val Ala
344          530          535          540
347 Gln Met Pro Leu Val Glu Ala Lys Leu Arg Phe Ile Gln Ala Trp Gln
348 545          550          555          560
351 Ser Leu Pro Glu Phe Gly Leu Thr Tyr Tyr Leu Val Arg Phe Lys Gly
352          565          570          575
355 Ser Lys Lys Asp Asp Ile Leu Gly Val Ser Tyr Asn Arg Leu Ile Lys
356          580          585          590
359 Ile Asp Ala Ala Thr Gly Ile Pro Val Thr Thr Trp Arg Phe Thr Asn
360          595          600          605
363 Ile Lys Gln Trp Asn Val Asn Trp Glu Thr Arg Gln Val Val Ile Glu
364          610          615          620
367 Phe Asp Gln Asn Val Phe Thr Ala Phe Thr Cys Leu Ser Ala Asp Cys
368 625          630          635          640
371 Lys Ile Val His Glu Tyr Ile Gly Gly Tyr Ile Phe Leu Ser Thr Arg
372          645          650          655
375 Ser Lys Asp Gln Asn Glu Thr Leu Asp Glu Asp Leu Phe His Lys Leu
376          660          665          670
379 Thr Gly Gly Gln Asp
380          675
383 <210> SEQ ID NO: 3
384 <211> LENGTH: 5
385 <212> TYPE: PRT
386 <213> ORGANISM: Unknown

```

**Please Note:**

Use of n and/ r Xaa have be n detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/733,757

DATE: 06/01/2001

TIME: 11:16:58

Input Set : A:\es.txt

Output Set: C:\CRF3\06012001\I733757.raw

L:399 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3